PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Glycated haemoglobin A1c as a risk factor of cardiovascular outcomes and all-cause mortality in diabetic and non-diabetic populations: a systematic review and meta-analysis.
AUTHORS	Cavero-Redondo, Iván; Peleteiro, Barbara; Álvarez-Bueno, Celia; Rodriguez-Artalejo, Fernando; Martinez-Vizcaino, Vicente

VERSION 1 - REVIEW

REVIEWER	Stevens, Sarah University of Oxford
REVIEW RETURNED	25-Jan-2017

GENERAL COMMENTS	My comments on the manuscript relate primarily to the selection criteria and statistical analysis. Firstly, it is unclear why studies that did not use Cox proportional hazards models were not included in the review (page 5, line 6) and this decision is not justified adequately. Since over a third of the studies excluded at full text stage were excluded for this reason (Figure 1), this presents a major limitation as valid estimates of the relative hazard can be obtained from other modelling strategies. It would perhaps be better to include papers reporting estimates of the relative hazard from any appropriate survival model and to examine the impact of the type of analysis (Cox PH or other) in sensitivity analysis. The meta-analysis method is appropriate, although the authors state that they pooled HR estimates (p5, line 56). Presumably they converted these onto the log-scale before pooling? A check on one set of results suggests this is the case, so the text should accurately reflect this. It is unclear why the authors failed to convert as many hazard ratios as possible onto a common scale, before they were pooled, rather than carrying out several different analyses for each reference level. For example, the hazard ratios in the analysis that is summarised in Fig 2a, 2b and 2d could all have been converted to reflect a common reference standard (HbA1c<6%). This would allow more data to pooled at once, and give more precise estimates. In all of the figures, it would be helpful to include the confidence intervals on the graphs and not just in the text. Finally, it is well known that Egger's test has low power for small numbers of studies (p6 line 35). The authors should consider calculating the number of null effect studies of mean weight that would need to be included in meta-analyses to result in a non-significant pooled effect (Rosenberg MS. The file-drawer problem

	revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. Evolution 2005;59:464-8). This analysis would be particularly pertinent given the high number of studies excluded because they did not use Cox PH models (see above point). More generally, the authors make reference to several previous reviews in this area. There needs to be greater clarity about the novelty and interest provided by this review.

REVIEWER	Jan Westerink University Medical Center Utrecht, The Netherlands
REVIEW RETURNED	21-Feb-2017

GENERAL COMMENTS	In this study the authors investigate the relation between HbA1c and risk of cardiovascular outcomes and mortality in a systematic review and meta-analysis of observational studies. They try to find optimal HbA1c levels. This article is based on systematic and comprehensive work, which again convincingly shows that not only the highest but also the lowest levels of HbA1c are detrimental. However, because of the design of this meta-analysis using a categorical HbA1c approach instead of a continuous approach as in earlier meta-analyses, the importance and novelty of this research on top of earlier meta-analyses might be limited. I have several concerns that need particular attention. Major concerns • Mortality is presented as if this is the main outcome measurement in this article. I have not been able to review the search strategy as this is not included. However, it seems no search terms have been included for “mortality”. Also articles are exluded in case of no information of cause of death, although these excluded articles in fact might have contributed to the meta-analyses for all-cause mortality. This could have contributed to the existing publication bias. Please clarify this issue. • The authors chose to use a categorical HbA1c approach. 19 articles are excluded because no data could be extracted to compare different levels of HbA1c. Probably these excluded observational studies had a continuous approach and presented a HR per 1SD or 1% increase in HbA1c level. The fact that these articles had to be removed should be noted as a limitation of this design which probably has further contributed to existence of the publication bias. • In the manuscript there is no reference to the flowchart. I do not understand why only 46 articles are used for the actual meta-analysis, whereas 74 articles are eligible. Please elaborate on this point and consider how this might have influenced the results. • No information is provided on the confounders included in the models of the included observational studies. • Risk of bias is properly assessed using the QUIPS tool. It would be interesting to repeat the meta-analyses in subgroups of low and low-moderate risk of bias articles. • Confidence intervals should be added to figures 2-5. • The knowledge gain due to this meta-analysis on top of previous meta-analyses should be addressed more extensively in the discussion section. • The existing publication bias for all comparisons is a serious limitation of this meta-analysis. Therefore it cannot be dismissed as

	being common to meta-analyses. In the discussion this limitation should be addressed properly with potential causes of this publication bias. • In the conclusion in the manuscript the optimal HbA1c ranges in diabetes and non-diabetes populations are interchanged. This should be corrected. Minor concerns • "Incidence of cardiovascular events" should be consequently changed to "risk of cardiovascular events" throughout the manuscript. • Instead of "predictor" as stated in the title or "prognostic biomarker" I would suggest to consistently use "risk factor". • Abstract: Consider to change "DerSimonian and Laird method" into "random effect models". This specific information might be better placed in the methods section of the manuscript • The search strategy is missing in the supplemental files.

VERSION 1 – AUTHOR RESPONSE

Response to reviewers
Reviewer: 1
Reviewer Name: S Stevens
My comments on the manuscript relate primarily to the selection criteria and statistical analysis. Firstly, it is unclear why studies that did not use Cox proportional hazards models were not included in the review (page 5, line 6) and this decision is not justified adequately. Since over a third of the studies excluded at full text stage were excluded for this reason (Figure 1), this presents a major limitation as valid estimates of the relative hazard can be obtained from other modelling strategies. It would perhaps be better to include papers reporting estimates of the relative hazard from any appropriate survival model and to examine the impact of the type of analysis (Cox PH or other) in sensitivity analysis.
Authors:
Considering that only one study reported an assessment or mortality risk including OR as effect measure, and using HbA1c categories comparable to ours, we have decided to include these studies in a supplementary table specifying the reference categories for each of them.
Moreover, we have included this in the Limitations section:
"Finally, of the six excluded studies reporting relative risk or odds ratio estimates, only one study reported an assessment of mortality risk including odds ratio as a measure of effect and using categories of HbA1c comparable to those of the other studies (Table F in the Supplementary File). Therefore, although the exclusion of these studies might bias our pooled estimates, these studies provided risk estimates similar to the pooled estimates obtained."

The meta-analysis method is appropriate, although the authors state that they pooled HR estimates (p5, line 56). Presumably they converted these onto the log-scale before pooling? A check on one set of results suggests this is the case, so the text should accurately reflect this.
Authors:
As suggested, we have included the information regarding this concern in the Statistical analysis and data synthesis section as follows:
"For each HR estimate, the lnHR was calculated by converting to the natural log scale."

It is unclear why the authors failed to convert as many hazard ratios as possible onto a common scale, before they were pooled, rather than carrying out several different analyses for each reference level. For example, the hazard ratios in the analysis that is summarised in Fig 2a, 2b and 2d could all

have been converted to reflect a common reference standard (HbA1c<6%). This would allow more data to pooled at once, and give more precise estimates. In all of the figures, it would be helpful to include the confidence intervals on the graphs and not just in the text.
Authors:
Thank you for the comments. As suggested, we have converted hazard ratios onto a common scale and included these data as supplementary material. Thus, we have included information regarding this new analysis in both the Statistical analysis and data synthesis and Results sections. We have also included confidence intervals on the graphs.
“Additionally, pooled HRs were calculated for each HbA1c level without comparison to any reference level to show the mortality estimates on a common scale.”
“When the pooled HR for each HbA1c level was calculated, independently of the HbA1c reference level, mortality risk increased at HbA1c levels <6.0%, 8.0% to 9.0% and >9.0% ($p < 0.001$)”
“When the pooled HR for each HbA1c level was calculated, independently of the HbA1c reference level, mortality risk increased at HbA1c levels <5.0% and >6.0% ($p < 0.05$).”

Finally, it is well known that Egger's test has low power for small numbers of studies (p6 line 35). The authors should consider calculating the number of null effect studies of mean weight that would need to be included in meta-analyses to result in a non-significant pooled effect (Rosenberg MS. The file-drawer problem revisited: a general weighted method for calculating fail-safe numbers in meta-analysis. Evolution 2005;59:464-8). This analysis would be particularly pertinent given the high number of studies excluded because they did not use Cox PH models (see above point).
Authors:
As the reviewer suggested, we have used the fail-safe N test, and we have included, in the Statistical analysis and data synthesis section, this information as follows:
“Also, Rosenthal's fail-safe N method was used to determine the number of unpublished or un-retrieved null studies that would be needed to increase the p-value above 0.05 (to make the effect not statistically significant)23”

Also in the Results section:
“Egger’s test showed potential publication bias for all-cause mortality in diabetics at HbA1c ranging from 6.0% to 7.0%, using HbA1c levels below 6.0% as the reference ($p = 0.006$; Fail-safe N test = 0), and at HbA1c below 6.0% using HbA1c ranging from 7.0 to 8.0% as the reference level ($p = 0.046$; Fail-safe N test = 197). For cardiovascular mortality in non-diabetics, potential publication bias was detected at HbA1c levels above 6.5% when HbA1c below 5.5% was the reference ($p = 0.029$; Fail-safe N test = 14) and at HbA1c below 5.0% when HbA1c ranging from 5.0% to 5.9% was the reference level ($p = 0.048$; Fail-safe N test = 1).”

More generally, the authors make reference to several previous reviews in this area. There needs to be greater clarity about the novelty and interest provided by this review.
Authors:
We have included in the Discussion section a paragraph supporting what is new in this review:
“Although previous studies have reported an increase in mortality risk by each 1% increase in HbA1c, their estimates are based on the assumption of a linear relationship between these variables, which the data from studies included in this review did not show. Thus, providing estimates of changes in mortality risk by HbA1c level categories did not presuppose any functional statistical relationship between the involved variables.”
In addition, in the Strengths and Limitations section:
“Previous meta-analyses have reported pooled estimates of the increase in mortality risk by each 1% increase in HbA1c; thus, their estimates are based on the assumption of a linear relationship between these variables, which the data from studies included in this review did not show. Our study provides pooled estimates of changes in mortality risk by HbA1c level categories, and therefore, did not presuppose any functional statistical relationship between the involved variables.”

Reviewer: 2
Reviewer Name: Jan Westerink
In this study the authors investigate the relation between HbA1c and risk of cardiovascular outcomes and mortality in a systematic review and meta-analysis of observational studies. They try to find optimal HbA1c levels. This article is based on systematic and comprehensive work, which again convincingly shows that not only the highest but also the lowest levels of HbA1c are detrimental. However, because of the design of this meta-analysis using a categorical HbA1c approach instead of a continuous approach as in earlier meta-analyses, the importance and novelty of this research on top of earlier meta-analyses might be limited. I have several concerns that need particular attention.
Authors:
Thank you for reviewer´s comments. As suggested by another reviewer, we have included in the Discussion section a paragraph supporting what is new in this review:
"Although previous studies have reported an increase in mortality risk by each 1% increase in HbA1c, their estimates are based on the assumption of a linear relationship between these variables, which the data from studies included in this review did not show. Thus, providing estimates of changes in mortality risk by HbA1c level categories did not presuppose any functional statistical relationship between the involved variables"
In addition, in the Strengths and Limitations section:
"Previous meta-analyses have reported pooled estimates of the increase in mortality risk by each 1% increase in HbA1c; thus, their estimates are based on the assumption of a linear relationship between these variables, which the data from studies included in this review did not show. Our study provides pooled estimates of changes in mortality risk by HbA1c level categories, and therefore, did not presuppose any functional statistical relationship between the involved variables."

Major concerns
• Mortality is presented as if this is the main outcome measurement in this article. I have not been able to review the search strategy as this is not included. However, it seems no search terms have been included for "mortality". Also articles are exluded in case of no information of cause of death, although these excluded articles in fact might have contributed to the meta-analyses for all-cause mortality. This could have contributed to the existing publication bias. Please clarify this issue.
Authors:
Several mortality-related terms had been included in the search strategy, such as "all-cause mortality", "cardiovascular mortality", "cause-specific mortality", death or "cardiovascular death" (please see Table A in Supplementary file). In addition, we have also included an example of the full search strategy in this supplementary material.

As suggested, we have modified exclusion criteria as follows:
"iii) studies not reporting risk of mortality or cardiovascular outcomes, such as myocardial infarction, stroke, major adverse cardiovascular events, coronary heart disease and heart failure"

• The authors chose to use a categorical HbA1c approach. 19 articles are excluded because no data could be extracted to compare different levels of HbA1c. Probably these excluded observational studies had a continuous approach and presented a HR per 1SD or 1% increase in HbA1c level. The fact that these articles had to be removed should be noted as a limitation of this design which probably has further contributed to existence of the publication bias.
Authors:
As suggested, we have included a sentence regarding this concern as a limitation in the Discussion section:
"Moreover, since 26 studies were excluded because they provided results using a continuous approach, and they presented an HR estimate corresponding to an increase of 1 SD or 1% in HbA1c level, this could contribute to publication bias. In order to assess the magnitude of this bias, we

determined the number of unpublished or unrecovered null studies that would have been published to make the effect not statistically significant using the Rosenthal fail-safe N method.23”

• In the manuscript there is no reference to the flowchart. I do not understand why only 46 articles are used for the actual meta-analysis, whereas 74 articles are eligible. Please elaborate on this point and consider how this might have influenced the results.
Authors:
We have included a reference to the flowchart in the Results section. Regarding the studies that have been included in the review, but not in the meta-analyses, this is because they provided results based on HbA1c categories not comparable with the other studies. We have included a sentence in the Limitations section regarding this concern as follows:
“Furthermore, from the 74 retrieved studies, only 46 were included in the pooled estimates, the other 28 studies reported their results using HbA1c categories not comparable with the included studies; thus, publication bias cannot be disregarded, although these excluded studies provided HR estimates similar to the pooled estimates obtained.”

• No information is provided on the confounders included in the models of the included observational studies.
Authors:
Thank you for reviewer´s comment. We have included, as supplementary material, a table detailing the covariates used for adjusting the data reported by the included studies. The following sentence was added to the Results section:
“The data extracted from the included studies were adjusted for several covariates (Table C in the Supplementary File).”

• Risk of bias is properly assessed using the QUIPS tool. It would be interesting to repeat the meta-analyses in subgroups of low and low-moderate risk of bias articles.
Authors:
As the reviewer suggests, we have conducted the meta-analysis based on risk of bias assessed by using the QUIPS tool. In addition, we have included, as supplementary material, a table with the results of the subgroup analyses.
Regarding this concern, we have included in the manuscript:
- A sentence in the Statistical analysis and data synthesis section as follows:
“Subgroup analyses were performed based on the risk of bias assessed by the QUIPS tool (low, moderate or high risk of bias).”

- In the Results section:
“Subgroup analyses based on the risk of bias assessed by the QUIPS tool showed a decrease in some pooled HR estimates when the analysis was performed in studies with low risk of bias (Table E in the Supplementary File).”

• Confidence intervals should be added to figures 2-5.

Authors:
Done. Thank you.

• The knowledge gain due to this meta-analysis on top of previous meta-analyses should be addressed more extensively in the discussion section.
Authors:
We have included in the Discussion section a paragraph supporting what is new in this review:
“Although previous studies have reported an increase in mortality risk by each 1% increase in HbA1c, their estimates are based on the assumption of a linear relationship between these variables, which

the data from studies included in this review did not show. Thus, providing estimates of changes in mortality risk by HbA1c level categories did not presuppose any functional statistical relationship between the involved variables."
In addition, in the Strengths and Limitations section:
"Previous meta-analyses have reported pooled estimates of the increase in mortality risk by each 1% increase in HbA1c; thus, their estimates are based on the assumption of a linear relationship between these variables, which the data from studies included in this review did not show. Our study provides pooled estimates of changes in mortality risk by HbA1c level categories, and therefore, did not presuppose any functional statistical relationship between the involved variables."

• The existing publication bias for all comparisons is a serious limitation of this meta-analysis. Therefore it cannot be dismissed as being common to meta-analyses. In the discussion this limitation should be addressed properly with potential causes of this publication bias.
Authors:
Thank you for the reviewer's comment. We have included a potential cause of publication bias.
"Among the reasons for this publication bias, we can highlight that due to the wide segmentation of data from the studies included in this review, some meta-analyses included a small number of studies, thus publication bias is likely. Furthermore, from the 74 retrieved studies, only 46 were included in the pooled estimates, the other 28 studies reported their results using HbA1c categories not comparable with the included studies; thus, publication bias cannot be disregarded, although these excluded studies provided HR estimates similar to the pooled estimates obtained. Moreover, since 26 studies were excluded because they provided results using a continuous approach, and they presented an HR estimate corresponding to an increase of 1 SD or 1% in HbA1c level, this could contribute to publication bias. In order to assess the magnitude of this bias, we determined the number of unpublished or unrecovered null studies that would have been published to make the effect not statistically significant using the Rosenthal fail-safe N method.23"

• In the conclusion in the manuscript the optimal HbA1c ranges in diabetes and non-diabetes populations are interchanged. This should be corrected.
Authors:
Thank you. Done.

Minor concerns
• "Incidence of cardiovascular events" should be consequently changed to "risk of cardiovascular events" throughout the manuscript.
Authors:
Thank you. Done.

• Instead of "predictor" as stated in the title or "prognostic biomarker" I would suggest to consistently use "risk factor".
Authors:
Thank you. Done.

• Abstract: Consider to change "DerSimonian and Laird method" into "random effect models". This specific information might be better placed in the methods section of the manuscript
Authors:
Thank you. Done.

• The search strategy is missing in the supplemental files.
Authors:
Thank you for the reviewer´s comment. We have included a table in the supplementary material with the search strategy for MEDLINE.

VERSION 2 – REVIEW

REVIEWER	Stevens, Sarah University of Oxford, UK
REVIEW RETURNED	26-Apr-2017

GENERAL COMMENTS	Abstract: The abstract should note that fewer than 74 studies were included in the meta-analyses. Statistical analysis: I am not sure that the authors fully understood my previous comment about converting all hazard ratios onto a common scale. Instead of presenting results according to the reference categories used in the original studies, results from all studies could be converted onto a common scale for risk in each category compared to (e.g.) HbA1c<6% before meta-analysis, hence producing one analysis rather than 4. The authors state that they have addressed this point and have added text to the manuscript but it is not clear from the methods section exactly what has been done and this needs clarification. I presume this extra analysis is also presented in Figure G in the supplementary material, but this is not cross-referenced in the text and Figure G does not show any reference categories as I would expect. The authors may want to note in the discussion that the use of categories in the source studies may have contributed to publication bias and bias in effect estimates in general because categorising continuous variables can inflate the type 1 error rate (See Royston et al, Statistics in medicine, 2006 (25) 1; 127-41 or Austin et al, Statistics in medicine, 2004 (23) 7; 1159-78)

REVIEWER	Jan Westerink UMC Utrecht, The Netherlands
REVIEW RETURNED	18-Apr-2017

GENERAL COMMENTS	The authors have addressed all comments adequately.

VERSION 2 – AUTHOR RESPONSE

Reviewer: 2
Reviewer Name: Jan Westerink

The authors have addressed all comments adequately.
Authors:
Thank you for the comment.

Reviewer: 1
Reviewer Name: S Stevens

Abstract: The abstract should note that fewer than 74 studies were included in the meta-analyses.
Authors:
“Seventy-four published studies were included in the systematic review, but only 46 studies could be incorporated in the meta-analysis…”

Statistical analysis: I am not sure that the authors fully understood my previous comment about converting all hazard ratios onto a common scale. Instead of presenting results according to the reference categories used in the original studies, results from all studies could be converted onto a common scale for risk in each category compared to (e.g.) HbA1c<6% before meta-analysis, hence producing one analysis rather than 4. The authors state that they have addressed this point and have added text to the manuscript but it is not clear from the methods section exactly what has been done and this needs clarification. I presume this extra analysis is also presented in Figure G in the supplementary material, but this is not cross-referenced in the text and Figure G does not show any reference categories as I would expect.
Authors:
Thank you for the comments. We have converted hazard ratios onto a common scale and included it as supplementary material. We have also included information regarding this new analysis in Statistical analysis and data synthesis and Results sections.
In Statistical analysis section:
"Additionally, for all-cause mortality, HR estimates for each HbA1c category from all studies were converted onto a common scale using HbA1c <6.0% as reference category for diabetic population and HbA1c <5.0% for non-diabetic population. HR estimate was reciprocated from risk or protective factor to reference value.21"

In Results section:
"When pooled HR was calculated converting HbA1c <6.0% as reference level, an increase of risk was shown at HbA1c level >9.0% ($p < 0.001$). Conversely, HbA1c levels 6.0% to 7.0% and 7.0% to 8.0% were presented as protective factor ($p < 0.001$) (Figure G in the Supplementary File)…

When pooled HR was calculated converting HbA1c <6.0% as reference level an increase of risk was showed at HbA1c level>6.0% ($p < 0.05$) (Figure G in the Supplementary File)."

The authors may want to note in the discussion that the use of categories in the source studies may have contributed to publication bias and bias in effect estimates in general because categorising continuous variables can inflate the type 1 error rate (See Royston et al, Statistics in medicine, 2006 (25) 1; 127-41 or Austin et al, Statistics in medicine, 2004 (23) 7; 1159-78)
Authors:
"Furthermore, using categories in the included studies may have contributed to publication bias and bias in HR estimates because categorizing continuous variables can increase the type 1 error rate.41"

VERSION 3 – REVIEW

REVIEWER	Stevens, Sarah University of Oxford, UK
REVIEW RETURNED	08-May-2017

GENERAL COMMENTS	All comments have been addressed.